A nonlinear canonical form for reduced order observer design

Driss Boutat, Gang Zheng, Hassan Hammouri

A nonlinear canonical form for reduced order observer design

Driss Boutat * **Gang Zheng** ** **Hassan Hammouri** ***

* *ENSI de Bourges, Insitut PRISME, 88 Boulevard de Lahitolle, 18020 Bourges, France (e-mail: driss.boutat@ensi-bourges.fr).*
** *INRIA Lille-Nord Europe, 40, avenue Halley, 59650 Villeneuve d'Ascq, France (e-mail: gang.zheng@inria.fr).*
*** *Université Lyon 1; CNRS UMR 5007 LAGEP; 43 bd du 11 novembre, 69100 Villeurbanne, France (e-mail: hammouri@lagep.cpe.fr).*

Abstract: This paper presents a nonlinear canonical form which is used for the design of a reduced order observer. Sufficient and necessary geometric conditions are given in order to transform a special class of nonlinear systems to the proposed nonlinear canonical form and the corresponding reduced order observer is analyzed.

Keywords: Canonical form, Reduced order observer

1. INTRODUCTION

The observer design for nonlinear dynamical systems is an important problem in the field of control theory. For nonlinear systems, several techniques, such as Luenberger observer Luenberger (1971), hight gain observer Hammouri and Gautier (1988) Gauthier et al. (1985), Busawon et al. (1998) and so on, are proposed to design nonlinear observers for different cases, and a general formalism to design nonlinear observer for generic nonlinear systems is still missing. Since the beginning of the 1980's, many significant researches were done on the problem of transforming nonlinear dynamical systems into simple normal observable forms, based on which one can apply existing observer techniques, from algebraic and geometric points of view, see Xia and Gao (1989), Respondek et al. (2004), Krener and Respondek (1985), Krener and Isidori (1983), Glumineau et al. (1996), Bestle and Zeitz (1983), Back et al. (2005), Boutat et al. (2009), Zheng et al. (2007), Zheng et al. (2009).

Roughly speaking, two types of observers can be classified. The first class is so-called full order observer which estimates all states of the system, including the measurable states which are the outputs. Obviously this redundant estimations of measurable states are not necessary, and conversely it might increase the complexity of the observer design and practical realization. That is the reason why the second class of observers was born, named as reduced order observer, which, different from full order observers, needs to estimate only unmeasurable states of the studied system. It was firstly introduced for linear systems by Luenberger (1971) to reduce the number of dynamical equations by estimating only the unmeasurable states. Necessary and sufficient conditions for the existence of minimal reduced order observer for linear systems were presented in Darouach (2000). Then it was generalized for nonlinear dynamical systems by imposing the Lipschitz conditions for nonlinear terms, see Xu (2009), and invariant manifold Karagiannis et al. (2008). The problem of designing full order observer and reduced order observer for nonlinear systems with linearizable error dynamics and its application to synchronization problem was analyzed in Nijmeijer and Mareels (1997), in which authors pointed out that the existence of a full order observer with linear error dynamics implies the existence of a reduced order observer with linear error dynamics, however the reverse is not valid. Moreover, there are no results available to provide conditions, under which via a transformation, a reduced order observer with linear error dynamics may be found.

In this paper, we give a new nonlinear canonical form which allows us to design reduced order observers, just like the linear case. And necessary and sufficient conditions are given to guarantee the existance of the proposed nonlinear canonical form.

2. A NONLINEAR CANONICAL FORM FOR REDUCED ORDER OBSERVER

Before giving the nonlinear canonical form which will be studied in this paper, let us give a definition of the so-called reduced order observer. Without loss of generality, consider the following nonlinear dynamical system:

$$\begin{cases} \dot{x} = F(x) \\ y = h(x) \end{cases} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state and $y \in \mathbb{R}^{m+p}$ is the output where $m + p < n$, $F : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^{m+p}$ are smooth vector functions. Without loss of generality, we assume that the components of outputs $h = (h_1, \cdots , h_{m+p})^T$ are linearly independent, and system (1) is observable.

By setting $x_2 = h(x) \in \mathbb{R}^{m+p}$, and choosing $n-m-p$ complementary variables:

$$x_1 = (x_{1,1}, \cdots, x_{1,n-m-p})^T$$

then system (1) can be decomposed into the following form:

$$\begin{cases} \dot{x}_1 = F_1(x) \\ \dot{x}_2 = F_2(x) \\ y = x_2 \end{cases} \tag{2}$$

where $F_1(x)$ and $F_2(x)$ are relatively determined by the choice of x_1 and the dynamics $F(x)$ defined in (1).

For (2), we try to design a reduced order observer to estimate x_1.

Definition 1. The dynamical system defined as follows:

$$\dot{\hat{x}}_1 = \widetilde{F}_1(\hat{x}_1, x_2)$$

where x_2 is the output of (2), is a symptomatically reduced order observer for (2) if

$$\lim_{t\to\infty} \parallel \hat{x}_1(t) - x_1(t) \parallel = 0.$$

Moreover, it is said to be an exponentially reduced order observer if

$$\parallel \hat{x}_1(t) - x_1(t) \parallel \leq a e^{-bt} \parallel \hat{x}_1(0) - x_1(0) \parallel$$

for $t > 0$, where a, b are both positive constants.

In what follows, we will present first the nonlinear canonical form which will be studied in this paper, then we will show that an exponentially reduced order observer can be easily designed for the proposed nonlinear canonical form. The sufficient and necessary geometric conditions to transform a generic nonlinear system to such an observable nonlinear canonical form will be detailed in the next section.

Let us consider the following observable nonlinear canonical form:

$$\begin{cases} \dot{z}_i = A_i z_i + \beta_i(y_1) z_o + \rho_i(y) \quad \text{for} \quad i = 1:m \\ \dot{\xi} = \alpha_1(y_1) z_o + \alpha_2(y) \\ \dot{\eta} = \mu(z, \xi, \eta) \\ y = \left(y_1^T, y_2^T\right)^T = \left(\xi^T, \eta^T\right)^T \end{cases} \tag{3}$$

where

$$\begin{aligned} z_i &= (z_{i,1}, \cdots, z_{i,r_i})^T \in \mathbb{R}^{r_i} \\ z_o &= (z_{1,r_1}, \cdots, z_{m,r_m})^T \in \mathbb{R}^m \\ \rho_i &= (\rho_{i,1}, \cdots, \rho_{i,r_i})^T \in \mathbb{R}^{r_i} \\ \xi &= (\xi_1, \cdots, \xi_m)^T \in \mathbb{R}^m \\ \eta &= (\eta_1, \cdots, \eta_p)^T \in \mathbb{R}^p \\ \alpha_2 &= (\alpha_{2,1}, \cdots, \alpha_{2,m})^T \in \mathbb{R}^m \end{aligned}$$

with $\sum_{i=1}^m r_i = n-m-p$ and the $r_i \times m$ matrix β_i, the $m \times m$ matrix α_1 and the $r_i \times r_i$ matrix A_i defined as follows:

$$A_i = \begin{pmatrix} 0 & \cdots & 0 & 0 & 0 \\ 1 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \cdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}$$

Remark 1. Since the nonlinear canonical form (3) is supposed to be observable, thus z_o in (3) can be observed from the output y_1, which implies $Rank(\alpha_1(y_1)) = m$.

In a more compact manner, system (3) can be rewritten as follows:

$$\dot{z} = Az + \beta(y_1) z_o + \rho(y) \tag{4}$$

$$\dot{\xi} = \alpha_1(y_1) z_o + \alpha_2(y) \tag{5}$$

$$\dot{\eta} = \mu(z, \xi, \eta) \tag{6}$$

$$y = \left(y_1^T, y_2^T\right)^T = \left(\xi^T, \eta^T\right)^T \tag{7}$$

where $A = diag[A_1, \cdots, A_m]$, $\beta = (\beta_1^T, \cdots, \beta_m^T)^T$, $\rho = (\rho_1^T, \cdots, \rho_m^T)^T$.

Let us denote C_i the $1 \times r_i$ vector defined as $C_i = [0, \cdots, 0, 1]$ for $1 \leq i \leq m$, then we can define a $m \times (n-m-p)$ matrix C as follows:

$$C = diag[C_1, \cdots, C_m] \tag{8}$$

which implies $z_o = Cz$.

For (4-7), if we can accurately measure y_1 and calculate $\dot{y}_1$, this allows us to define a "new" output Y being a function of known output y and the derivative of y_1 in (4-7):

$$Y = \alpha_1^{-1}(y_1)\left(\dot{y}_1 - \alpha_2(y)\right) \tag{9}$$

Then we can state the following result.

Proposition 1. The following dynamical system:

$$\dot{\hat{z}} = A\hat{z} + \beta(y_1) C\hat{z} + \rho(y) - K(y_1)(Y - C\hat{z}) \tag{10}$$

where $K(y_1) = -\beta(y_1) + \kappa$ and Y is defined in (9), is an exponentially reduced order observer for (4-7), if the chosen $(n-m-p) \times m$ matrix κ makes $(A + \kappa C)$ Hurwitz, where C is defined in (8).

Proof 1. Let $e = \hat{z} - z$ be the estimation error. Since $z_o = Cz$, then we can easily derive the dynamic of observation error from (4) and (10) as follows

$$\dot{e} = [A + (\beta(y_1) + K(y_1))C]e \tag{11}$$

Since the gain matrix $K(y_1)$ can be freely chosen, hence without loss of generality we set

$$K(y_1) = -\beta(y_1) + \kappa$$

which makes (11) become

$$\dot{e} = (A + \kappa C)e. \tag{12}$$

Consequently, if κ is chosen in such a way that matrix $(A + \kappa C)$ is Hurwitz, then the exponential convergence of $\hat{z}$ to z can be guaranteed.

Let us remark that the proposed reduced order observer (10) is based on the "new" output Y defined in (9), which clearly shows that the derivative of the real output y_1 should be calculated according to (9). However, it is well-known that the derivative of noisy signal should be avoided if possible in practice, since derivative operation will amplify the influence of noise. Hence, several techniques can be used to limit the influence of noise when computing the derivative of noisy signal. The most used way is to pass y_1 firstly through a low-pass filter and then calculate the derivative of y_1. It is also possible to calculate the derivative by algebraic method recently proposed in Fliess (2006), Fliess and Sira-Ramirez (2004), by converting the calculation of derivative to the calculation of integration, which is useful to annihilate noise.

In the following, a more practical observer is proposed based on the following hypothesis.

Hypothesis 1. We assume that the term $K(y_1)\alpha_1^{-1}(y_1)$ is integrable with respect to y_1, i.e. we can find a $\Gamma(y_1)$ such that

$$\frac{\partial \Gamma(y_1)}{\partial y_1} = K(y_1)\alpha_1^{-1}(y_1)$$

Based on Hypothesis 1, we can define ς as follows:

$$\varsigma = \hat{z} + \Gamma(y_1) \tag{13}$$

It should be noted that $\Gamma(y_1)$ defined in Hypothesis 1 can be considered as a signal of filtered y_1, and thus limit the influence of noise on y_1.

Inserting Y defined in (9) into (10), then (10) can be rewritten into

$$\begin{aligned}\dot{\hat{z}} + K(y_1)\alpha_1^{-1}(y_1)\dot{y}_1 &= (A + \beta(y_1)C + K(y_1)C)\hat{z} + \rho(y) \\ &\quad + K(y_1)\alpha_1^{-1}(y_1)\alpha_2(y)\end{aligned} \tag{14}$$

In order to avoid the derivative of y_1, we take the new variable ς into account, then a more practical reduced order observer can be derived from (14) as follows

$$\begin{aligned}\dot{\varsigma} &= (A + \beta(y_1)C + K(y_1)C)\left(\varsigma - \Gamma(y_1)\right) \\ &\quad + \rho(y) + K(y_1)\alpha_1^{-1}(y_1)\alpha_2(y_1) \\ &= (A + \kappa C)\varsigma + \rho(y) - (A + \kappa C)\Gamma(y_1) \\ &\quad + K(y_1)\alpha_1^{-1}(y_1)\alpha_2(y)\end{aligned} \tag{15}$$

with $\Gamma(y_1)$ defined in (13).

Remark 2. The new and more practical reduced order observer defined in (15) aims at limiting the influence of noise on the output, and it is based on Hypothesis 1. The estimation of z can be computed according to (13).

Corollary 1. Concerning the nonlinear canonical form with known inputs $u \in \mathbb{R}^q$ as follows:

$$\dot{z} = Az + \beta(y_1)z_o + \rho(y) + \epsilon(y, u) \tag{16}$$

$$\dot{\xi} = \alpha_1(y_1)z_o + \alpha_2(y) \tag{17}$$

$$\dot{\eta} = \mu\left(z, \xi, \eta\right) \tag{18}$$

$$y = \left(y_1^T, y_2^T\right)^T = \left(\xi^T, \eta^T\right)^T \tag{19}$$

an exponentially reduced order observer can be designed of the form:

$$\dot{\hat{z}} = A\hat{z} + \beta(y_1)C\hat{z} + \rho(y) + \epsilon(y, u) - K(y_1)(Y - C\hat{z})$$

with Y is defined in (9), and $K(y_1) = -\beta(y_1) + \kappa$ where κ is chosen in such a way that $(A + \kappa C)$ is Hurwitz with C defined in (8). Moreover, following the same procedure of deriving (15), a more practical reduced order observer can be deduced as well.

Example 1. Let us consider the following nonlinear system which is already of the form (3) as follows

$$\begin{cases} \dot{z}_{1,1} = y_1 z_{1,2} + z_{2,1}, \dot{z}_{1,2} = z_{1,1} + \rho_{1,2}(y) \\ \dot{z}_{2,1} = y_1^2 z_{1,2} \\ \dot{\xi}_1 = z_{1,2}, \dot{\xi}_2 = z_{2,1} \\ \dot{\eta} = \mu(z, \xi) \\ y_1 = (\xi_1, \xi_2)^T \\ y_2 = \eta \end{cases} \tag{20}$$

then one has

$$z_o = \begin{pmatrix} z_{1,2} \\ z_{2,1} \end{pmatrix}, C_1 = (0, 1), C_2 = 1, C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \beta(y_1) = \begin{pmatrix} y_1 & 1 \\ 0 & 0 \\ y_1^2 & 0 \end{pmatrix}$$

It can be checked that if one takes $\kappa = \begin{pmatrix} -8, & -3 \\ -4, & -2 \\ -2, & -1 \end{pmatrix}$, then $(A + \kappa C)$ is Hurwitz.

Hence, $K(y_1) = \begin{pmatrix} -y_1 - 8, & -4 \\ -4, & -2 \\ -y_1^2 - 2, & -1 \end{pmatrix}$ and it is easy to get $\Gamma(y_1) = \begin{pmatrix} \dfrac{-y_1^2}{2} - 8y_1, & -4y_1 \\ -4y_1, & -2y_1 \\ \dfrac{-y_1^3}{3} - 2y_1, & -y_1 \end{pmatrix}$. Consquently Hypothesis 1 is satisfied and we can design a more practical reduced order observer in the form of (15) for (20).

3. TRANSFORMATION TO THE NONLINEAR CANONICAL FORM

In the last section, we defined a new nonlinear canonical form, and its associated reduced order observer is discussed as well. This section is devoted to deducing necessary and sufficient geometric conditions which allows us to transform a nonlinear system into the proposed nonlinear canonical form (3).

Let us consider a class of nonlinear systems, where the generic nonlinear system (2) can be decomposed into the following form:

$$\dot{x} = F_1(x, \zeta, \vartheta) = f(x, \zeta, \vartheta) \tag{21}$$

$$\dot{\zeta} = F_{21}(x, \zeta, \vartheta) = \gamma_1(\zeta)H(x) + \gamma_2(\zeta, \vartheta) \tag{22}$$

$$\dot{\vartheta} = F_{22}(x, \zeta, \vartheta) = \varepsilon\left(x, \zeta, \vartheta\right) \tag{23}$$

$$y = \left(\zeta^T, \vartheta^T\right)^T \tag{24}$$

where $x \in \mathbb{R}^{n-m-p}$, $\zeta \in \mathbb{R}^m$, $\vartheta \in \mathbb{R}^p$, $y \in \mathbb{R}^{m+p}$, $f : \mathbb{R}^n \to \mathbb{R}^{n-m-p}$, γ_1 and γ_2 are appropriate dimensional smooth vector functions, $H(x) = (H_1(x), \cdots, H_m(x))^T$ are linearly independent. Moreover, we suppose system (21-24) is observable, and the observability indices, (see Krener and Respondek (1985), Marino and Tomei (1995)) for $H(x) = (H_1(x), \cdots, H_m(x))^T$ with respect to (21) are respectively noted as $(r_1, \cdots, r_m)$, such that[1] $r_1 \geq r_2 \geq \cdots \geq r_m \geq 1$ and $\sum_{i=1}^m r_i = n - m - p$.

Let us denote $\theta_{i,j} = dL_f^{j-1}H_i$ for $1 \leq i \leq m$ and $1 \leq j \leq r_i$. Because of the observability of system (21-24), then the codistrubition $span\{\theta_{i,j}, \text{ for } 1 \leq i \leq m, 1 \leq j \leq r_i\}$ is of rank $n - m - p$.

Let us note $F = \left(f^T, F_2^T\right)^T$ and $F_2 = \left(F_{21}^T, F_{22}^T\right)^T$, where the decomposition of F_2 into F_{21} and F_{22} makes the rank of $\gamma_1(\zeta)$ equal to m.

For $1 \leq i \leq m$, let us denote $\tau_{i,1}$ vectors determined by the following equations:

$$\begin{array}{ll} \theta_{i,r_i}(\tau_{i,1}) = 1, & \text{for } 1 \leq i \leq m \\ \theta_{i,k}(\tau_{i,1}) = 0, & \text{for } 1 \leq k \leq r_i - 1 \\ \theta_{j,k}(\tau_{i,1}) = 0, & \text{for } 1 \leq j < i \text{ and } 1 \leq k \leq r_i \\ \theta_{j,k}(\tau_{i,1}) = 0, & \text{for } i < j \leq m \text{ and } 1 \leq k \leq r_j \end{array} \tag{25}$$

Then, by induction we can define the following family of vector fields from $\tau_{i,1}$ as follows

$$\tau_{i,j} = [\tau_{i,j-1}, f] \text{ for } 1 \leq i \leq m \text{ and } 2 \leq j \leq r_i$$

As we will prove in the following theorem that a necessary condition to transform (21-24) into (4-7) is

$$[\tau_{i,j}, \tau_{s,l}] = 0$$

[1] It is possible by reordering H_i for $1 \leq i \leq m$.

for $1 \leq i \leq m$, $1 \leq j \leq r_i$, $1 \leq s \leq m$ and $1 \leq l \leq r_s$. Suppose that this condition is satisfied, then we can construct m vector fields $\sigma_1, \cdots, \sigma_m$ and p vector fields $\upsilon_1, \cdots, \upsilon_p$ such that $\{\tau_{i,j}, \sigma_k, \upsilon_l\}$ forms a basis for $1 \leq i \leq m$, $1 \leq j \leq r_i$, $1 \leq k \leq m$ and $1 \leq l \leq p$, by the following equations:

$$\begin{array}{l} d\zeta_i(\sigma_k) = \delta_i^k,\ d\zeta_i(\upsilon_l) = 0,\ \text{for } 1 \leq i \leq m,\ 1 \leq k \leq m,\ 1 \leq l \leq p \\ d\vartheta_l(\sigma_k) = 0,\ d\vartheta_l(\upsilon_s) = \delta_l^s,\ \text{for } 1 \leq j \leq r_i,\ 1 \leq l \leq p,\ 1 \leq s \leq p \end{array} \tag{26}$$

and

$$[\tau_{i,j}, \sigma_k] = [\sigma_s, \sigma_k] = [\tau_{i,j}, \upsilon_l] = [\upsilon_l, \upsilon_t] = [\sigma_s, \upsilon_l] = 0 \tag{27}$$

where δ_i^k represents Kronecker delta, i.e. $\delta_i^k = 1$ if $i = k$, otherwise $\delta_i^k = 0$.

Let us note

$$\begin{aligned} \theta = (&\theta_{1,1}, \cdots, \theta_{1,r_1}, \cdots, \theta_{m,1}, \cdots, \theta_{m,r_m}, \\ &d\zeta_1, \cdots, d\zeta_m, d\vartheta_1, \cdots, d\vartheta_p)^T \end{aligned}$$

and

$$\begin{aligned} \tau = (&\tau_{1,1}, \cdots, \tau_{1,r_1}, \cdots, \tau_{m,1}, \cdots, \tau_{m,r_m}, \\ &\sigma_1, \cdots, \sigma_m, \upsilon_1, \cdots, \upsilon_p) \end{aligned}$$

Set $\Lambda = \theta(\tau)$. Due to the observability rank condition, this matrix is invertible, hence we can define the following multi 1-forms

$$\omega = \Lambda^{-1}\theta = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} \tag{28}$$

where $\omega_2 = (d\zeta_1, \cdots, d\zeta_m, d\vartheta_1, \cdots, d\vartheta_p)^T$ and ω_1 is the rest of ω.

Now we are ready to claim our main result.

Theorem 1. There exists a diffeomorphism $(z, \xi, \eta) = \phi(x, \zeta, \vartheta)$ which transforms the dynamical system (21-24) into the nonlinear canonical form (4-7) if and only if the following conditions are satisfied

(1) $[\tau_{i,j}, \tau_{s,l}] = 0$, for $1 \leq i \leq m$, $1 \leq j \leq r_i$, $1 \leq s \leq m$ and $1 \leq l \leq r_s$;
(2) $\theta_{j,1}(\tau_{i,k}) = 0$, for $j > i$, $1 \leq i \leq m$, $r_j + 1 \leq k \leq r_i$;
(3) $[\tau_{i,r_i}, \bar{F}] = \sum_{l=1}^{m} \sum_{j=1}^{r_l} V_{l,j}(y_1)\tau_{l,j} + \sum_{k=1}^{m} W_k(y_1)\sigma_k$, for $1 \leq i \leq m$, $1 \leq j \leq r_i$ and $1 \leq k \leq p$;
(4) $[\tau_{i,j}, F_2] \in \ker \omega_1$ for $1 \leq i \leq m$ and $1 \leq j \leq r_i - 1$;

where $\bar{F} = \begin{pmatrix} f \\ F_{21} \end{pmatrix}$, $V_{i,j}(y_1)$ and $W_k(y_1)$ are smooth functions of y_1 defined in (24), σ_k is defined by (26-27) and ω_1 is defined in (28).

Proof 2. **Necessity:** Indeed, if (21-24) can be transformed into (4-7) via the diffeomorphism $(z, \xi, \eta) = \phi(x, \zeta, \vartheta)$, then $\tau_{i,j} = \frac{\partial}{\partial z_{i,j}}$ for $1 \leq i \leq m$, $1 \leq j \leq r_i$, $\sigma_k = \frac{\partial}{\partial \xi_k}$ for $1 \leq k \leq m$ and $\upsilon_l = \frac{\partial}{\partial \eta_l}$ for $1 \leq l \leq p$. And it is easy to check that all conditions of Theorem 1 are satisfied.

Sufficiency: Consider the multi 1-forms ω defined in (28), we have $\omega(\tau) = \mathbf{I}_{(n-m)\times(n-m)}$, which implies $\omega(\tau_{i,j})$, $\omega(\sigma_k)$ and $\omega(\upsilon_l)$ are constant. Therefore,

$$\begin{aligned} d\omega(\tau_{i,j}, \tau_{k,s}) &= L_{\tau_{i,j}}\omega(\tau_{k,s}) - L_{\tau_{k,s}}\omega(\tau_{i,j}) - \omega([\tau_{i,j}, \tau_{k,s}]) \\ &= -\omega([\tau_{i,j}, \tau_{k,s}]) \end{aligned}$$

thus, we can calculate m vector fields $\sigma_1, \cdots, \sigma_m$ and p vector fields $\upsilon_1, \cdots, \upsilon_p$, such that $\{\tau_{i,j}, \sigma_k, \upsilon_l\}$ forms a basis, satisfying (26) and (27). Following the same principle, we have

$$\begin{array}{l} d\omega(\tau_{i,j}, \sigma_k) = -\omega([\tau_{i,j}, \sigma_k]), d\omega(\tau_{i,j}, \upsilon_l) = -\omega([\tau_{i,j}, \upsilon_l]) \\ d\omega(\sigma_s, \sigma_k) = -\omega([\sigma_l, \sigma_k]), d\omega(\upsilon_l, \upsilon_t) = -\omega([\upsilon_l, \upsilon_t]) \\ d\omega(\sigma_k, \upsilon_l) = -\omega([\sigma_k, \upsilon_l]) \end{array}$$

Since ω is an isomorphism, this implies the equivalence between $[\tau_{i,j}, \tau_{k,s}] = 0$ and $d\omega = 0$.

According to theorem of Poincaré (see Abraham and Marsden (1966)), $d\omega = 0$ implies that there exists a local diffeomorphism $(z, \xi, \eta) = \phi(x, \zeta, \vartheta)$ such that $\omega = d\phi$. We note $\omega_i = d\phi_i$ for $1 \leq i \leq 2$.

Since condition (1) in Theorem 1 is satisfied, it implies $\phi_*(\tau_{i,j}) = \frac{\partial}{\partial z_{i,j}}$, $\phi_*(\sigma_k) = \frac{\partial}{\partial \xi_k}$ and $\phi_*(\upsilon_l) = \frac{\partial}{\partial \eta_l}$ for $1 \leq i \leq m$, $1 \leq j \leq r_i$, $1 \leq k \leq m$ and $1 \leq l \leq p$.

Now let us clarify the affect of this transformation on $f(x, \zeta, \vartheta)$ defined in (21). By the diffeomorphism $\phi(x, \zeta, \vartheta)$, we have $\begin{pmatrix} \dot{z} \\ \dot{\xi} \\ \dot{\eta} \end{pmatrix} = \phi_*(F)$ where $F = \begin{pmatrix} f \\ F_2 \end{pmatrix}$. It is easy to see that

$$\omega \begin{pmatrix} f \\ F_2 \end{pmatrix} = \begin{pmatrix} \omega_1(f + F_2) \\ \omega_2(f + F_2) \end{pmatrix} = \begin{pmatrix} \omega_1(f) + \omega_1(F_2) \\ \omega_2(f) + \omega_2(F_2) \end{pmatrix}$$

Then, for $1 \leq i \leq m$ and $1 \leq j \leq r_i - 1$, we get

$$\begin{aligned} \frac{\partial(\phi_*(F))}{\partial z_{i,j}} &= \begin{pmatrix} [\omega_1(\tau_{i,j}), \omega_1(f) + \omega_1(F_2)] \\ [\omega_2(\tau_{i,j}), \omega_2(f) + \omega_2(F_2)] \end{pmatrix} \\ &= \begin{pmatrix} \omega_1[\tau_{i,j}, f] + \omega_1[\tau_{i,j}, F_2] \\ [\omega_2\tau_{i,j}, \omega_2(f) + \omega_2(F_2)] \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{\partial}{\partial z_{i,j+1}} \\ [\omega_2(\tau_{i,j}), \omega_2(f) + \omega_2(F_2)] \end{pmatrix} \end{aligned}$$

since condition (4) implies $\omega_1[\tau_{i,j}, F_2] = 0$.

By integrating we obtain:

$$\omega_1(F) = Az + \varrho(y, z_o)$$

Then, let us prove that the diffeomorphism $\phi(x, \zeta, \vartheta)$ will transform (22) to (5).

As we know

$$\frac{\partial}{\partial z_{i,k}} H_j \circ \phi^{-1} = dH_j(\tau_{i,k}) = \theta_{j,1}(\tau_{i,k})$$

According to the definition of $\tau_{i,1}$ in (25), we get

$$\begin{aligned} \theta_{j,1}(\tau_{i,k}) &= \theta_{j,1}([\tau_{i,k-1}, f]) = \theta_{j,2}(\tau_{i,k-1}) \\ &= \cdots = \theta_{j,k}(\tau_{i,1}) = 0 \end{aligned}$$

for $j < i$ and $1 \leq k \leq r_i$. Following the same procedure, we have $\theta_{j,1}(\tau_{i,k}) = 0$, for $j > i$ and $1 \leq k \leq r_j$. Combined with condition (2) in Theorem 1, we have

$$\theta_{j,1}(\tau_{i,k}) = \begin{cases} 1,\ i = j, k = r_i \\ 0,\ \text{otherwise} \end{cases}$$

which implies (22) can be written as

$$\dot{\zeta} = \gamma_1(\zeta)z_o + \gamma_2(\zeta) \tag{29}$$

via $(z, \xi, \eta) = \phi(x, \zeta, \vartheta)$. Hence, by setting $\omega_2 = d\phi_2$ where $\phi_2 = \mathbf{I}_{(m+p)\times(m+p)}$, then we get

$$\phi_*(F) = \begin{pmatrix} Az + \varrho(y, z_o) \\ \gamma_1(y_1)z_o + \gamma_2(y) \\ \mu(z, \xi, \eta) \end{pmatrix} \tag{30}$$

Finally, by the condition (3), we have

$$\begin{aligned}\frac{\partial\phi_*\left(\bar{F}\right)}{\partial z_{i,r_i}} &= \phi_*\left(\left[\tau_{i,r_i},\bar{F}\right]\right)\\ &= \phi_*\left(\sum_{l=1}^{m}\sum_{j=1}^{r_l}V_{l,j}(y_1)\tau_{l,j}+\sum_{k=1}^{m}W_k(y_1)\sigma_k\right)\\ &= \sum_{l=1}^{m}\sum_{j=1}^{r_l}V_{l,j}(y_1)\frac{\partial}{\partial z_{l,j}}+\sum_{k=1}^{m}W_k(y_1)\frac{\partial}{\partial \xi_k}\end{aligned}$$

which means that $\varrho(y,z_o)$ in (30) can be decomposed as:

$$\varrho(y,z_o)=\beta(y_1)z_o+\rho(y)$$

Thus we proved that (21-24) can be transformed to form (4-7) via ϕ.

Remark 3. As explained in the above proof, conditions (1), (2) and (4) of Theorem 1 are used to determine the diffeomorphism $(z,\xi,\eta)=\phi(x,\zeta,\vartheta)$, which transforms (21-22) to the form:

$$\begin{aligned}\dot{z} &= Az+\varrho(y,z_o)\\ \dot{\xi} &= \alpha_1(y_1)z_o+\alpha_2(y)\\ \dot{\eta} &= \mu\left(z,\xi,\eta\right)\end{aligned}$$

Condition (3) guarantees that the above form can be written in (4), i.e. $\varrho(y,z_o)=\beta(y_1)z_o+\rho(y)$.

Remark 4. If system (21) is with inputs $u\in\mathbb{R}^q$ of the following form

$$\dot{x}=f(x,\zeta,\vartheta)+\sum_{i=1}^{q}g_i(x,\zeta,\vartheta)u_i \tag{31}$$

then it can be transformed into (16), if Theorem 1 is valid and also the following condition is satisfied:

$$[\tau_{i,j},g_k]=0$$

for $1\le i\le m$, $1\le j\le r_i$ and $1\le k\le q$.

The reason is that, if $[\tau_{i,j},g_k]=0$, then

$$\frac{\partial}{\partial z_{i,j}}\phi_*(g_k)=\phi_*([\tau_{i,j},g_k])=0$$

which implies $\phi_*(g)=\nu(y)$, and thus (31) is transformed into (16).

Remark 5. In the case where $m=1$ for (21-24), there exists a diffeomorphism $(z,\xi,\eta)=\phi(x,\zeta,\vartheta)$ which transforms (21) in the following canonical form:

$$\begin{aligned}\dot{z} &= Az+\beta(y_1)+\rho(y)\\ \dot{\xi} &= \alpha_1(y_1)z_o+\alpha_2(y)\\ \dot{\eta} &= \mu\left(z,\xi,\eta\right)\\ y &= \left(\xi^T,\eta^T\right)^T\end{aligned}$$

which is an extension of the linear canonical form modulo an injection output studied in Krener and Isidori (1983). An example in this form can be found in Nijmeijer and Mareels (1997).

Corollary 2. There exists a diffeomorphism $(z,\xi,\eta)=\phi(x,\zeta,\vartheta)$ which transforms the dynamical system (21-24) into

$$\dot{z}=Az+\beta(y_1)z_o+\rho(y) \tag{32}$$

$$\dot{\xi}=\alpha_1(y_1)B\left(y\right)z_o+\alpha_2(y) \tag{33}$$

$$\dot{\eta}=\mu\left(z,\xi,\eta\right) \tag{34}$$

$$y=\left(\xi^T,\eta^T\right)^T \tag{35}$$

where A, β, γ are defined in (4-7) with

$$B\left(y\right)=\begin{pmatrix}1 & 0 & \cdots & 0 & 0\\ b_{2,1}(y) & 1 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ b_{m-1,1}(y) & b_{m-1,2}(y) & \cdots & 1 & 0\\ b_{m,1}(y) & b_{m,2}(y) & \cdots & b_{m,m-1}(y) & 1\end{pmatrix} \tag{36}$$

where $b_{i,j}(y)$ is a function of y for $2\le i\le m$ and $1\le j\le i-1$, if and only if the following conditions are satisfied

(1) Conditions (1), (3) and (4) of Theorem 1 are satisfied;
(2)

$$\theta_{j,1}(\tau_{i,k})=\begin{cases}0, & \text{for } j>i, 1\le i\le m, r_j+1\le k\le r_i-1\\ b_{j,i}(y), & \text{for } j>i, 1\le i\le m, k=r_i\end{cases}$$

where $b_{j,i}(y)$ is a function of y.

Proof 3. The proof of Corollary 2 follows the same argument as that of Theorem 1. Hence we only explain Condition (2) of Corollary 2.

Indeed, Condition (2) of Corollary 2 can be interpreted as, for $j>i$,

$$\frac{\partial}{\partial z_{i,r_i}}H_j\circ\phi^{-1}=\theta_{j,1}(\tau_{i,r_i})=b_{j,i}\left(y\right)$$

and $\frac{\partial}{\partial z_{i,k}}H_j\circ\phi^{-1}=0$ for $j>i$, $1\le i\le m$ and $r_j+1\le k\le r_i-1$. This, with the definition of $\tau_{i,1}$ in (25), yields

$$\begin{aligned}\frac{\partial}{\partial z_o}H\circ\phi^{-1} &= \begin{pmatrix}1 & 0 & \cdots & 0 & 0\\ b_{2,1}\left(y\right) & 1 & \cdots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ b_{m-1,1}\left(y\right) & b_{m-1,2}\left(y\right) & \cdots & 1 & 0\\ b_{m,1}\left(y\right) & b_{m,2}\left(y\right) & \cdots & b_{m,m-1}\left(y\right) & 1\end{pmatrix}\\ &= B\left(y\right)\end{aligned}$$

and $\frac{\partial}{\partial z_{i,k}}H\circ\phi^{-1}=0$ for $k\ne r_i$. By integration, we can prove that the diffeomorphism ϕ transforms (22) to (33) with an invertible matrix $B\left(y\right)$ defined in (36).

Remark 6. Since matrix $B\left(y\right)$ defined in (36) is invertible, the proposed observer of the form (10) is still valid, where the "new" output Y should be redefined as follows

$$Y=B^{-1}\left(y\right)\alpha^{-1}(y_1)\left(\dot{y}_1-\alpha_2(y)\right)$$

The following example highlights the validity of the proposed results.

Example 2. Let us consider the following nonlinear system

$$\begin{cases}\dot{x}_1=x_5x_1+x_3x_2, \dot{x}_2=x_1\\ \dot{x}_3=x_1, \dot{x}_4=x_2, \dot{x}_5=x_3\\ \dot{x_6}=x_3^2+x_2x_4+x_4x_5x_6\\ y_1=x_4, y_2=x_5, y_3=x_6\end{cases} \tag{37}$$

By setting $x=\left(x_1,x_2,x_3\right)^T$, $\zeta=\left(x_4,x_5\right)^T$ and $\vartheta=x_6$, we can rewrite (37) as follows

$$\begin{cases}\dot{x}=\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\\ \dot{x}_3\end{pmatrix}=\begin{pmatrix}\zeta_2x_1+x_3x_2\\ x_1\\ x_1\end{pmatrix}\\ \dot{\zeta}=\begin{pmatrix}\dot{\zeta}_1\\ \dot{\zeta}_2\end{pmatrix}=\begin{pmatrix}\dot{x}_4\\ \dot{x}_5\end{pmatrix}=\begin{pmatrix}x_2\\ x_3\end{pmatrix}\\ \dot{\vartheta}=x_3^2+x_2\zeta_1+\zeta_1\zeta_2\vartheta\\ y_1=\left(\zeta_1,\zeta_2\right)^T, y_2=\vartheta\end{cases} \tag{38}$$

which is of the form (4-7) with $H(x) = (x_2, x_3)^T$,$f = (\zeta_2 x_1 + x_3 x_2, x_1, x_1)^T$ $F_2 = \left(x_2, x_3, x_3^2 + x_2\zeta_1 + \zeta_1\zeta_2\vartheta\right)^T$ and $\bar{F} = (f, F_{21})^T = (\zeta_2 x_1 + x_3 x_2, x_1, x_1, x_2, x_3)^T$. Then we can define the following 1-forms:

$$\theta_{1,1} = dx_2,\ \theta_{1,2} = dx_1 \text{ and } \theta_{2,1} = dx_3$$
$$d\zeta_1 = dx_4,\ d\zeta_2 = dx_5 \text{ and } d\vartheta_1 = dx_6$$

According to (25), (26) and (27), we obtain

$$\tau_{1,1} = \frac{\partial}{\partial x_1}, \tau_{1,2} = \frac{\partial}{\partial x_2} + \frac{\partial}{\partial x_3} + x_5\frac{\partial}{\partial x_1}, \tau_{2,1} = \frac{\partial}{\partial x_3}$$
$$\sigma_1 = \frac{\partial}{\partial x_4}, \sigma_2 = \frac{\partial}{\partial x_5} + x_2\frac{\partial}{\partial x_1}, \upsilon_1 = \frac{\partial}{\partial x_6}$$

It is easy to check that $[\tau_{i,j}, \tau_{k,s}] = 0$ for $1 \le i \le 2$, $1 \le j \le r_i$, $1 \le s \le 2$ and $1 \le l \le r_s$ with $r_1 = 2$ and $r_2 = 1$. Moreover we have

$$[\tau_{1,2}, \bar{F}] = \zeta_2\tau_{1,2} + \sigma_2 + \sigma_1$$
$$[\tau_{2,1}, \bar{F}] = \sigma_2$$

and

$$\theta_{2,1}(\tau_{1,2}) = 1$$

In order to calculate the diffeomorphism, let us consider $\Lambda = \theta\tau$ which gives $\omega_1 = \begin{pmatrix} dx_1 - x_5 dx_2 - x_2 dx_5 \\ dx_2 \\ -dx_2 + dx_3 \end{pmatrix}$. Then $\omega_1[\tau_{1,1}, F_2] = 0$, implying $[\tau_{1,1}, F_2] \in \ker\omega_1$. Thus all conditions of Corollary 2 are satisfied, and we have

$$(z_{1,1}, z_{1,2}, z_{2,1}, \xi_1, \xi_2, \eta)^T = (x_1 - x_2 x_5, x_2, x_3 - x_2, x_4, x_5, x_6)^T$$

which transforms (37) into

$$\begin{cases} \dot{z}_{1,1} = 0, \dot{z}_{1,2} = z_{1,1} + \xi_2 z_{1,2}, \dot{z}_{2,1} = 0 \\ \dot{\xi}_1 = z_{1,2}, \dot{\xi}_2 = z_{2,1} + z_{1,2} \\ \dot{\eta} = (z_{1,2} + z_{2,1})^2 + z_{1,2}\xi_1 + \xi_1\xi_2\eta \\ y_1 = (\xi_1, \xi_2)^T, y_2 = \eta \end{cases}$$

4. CONCLUSION

A nonlinear canonical form was studied in this paper. We firstly gave a set of sufficient and necessary geometric conditions which transform a special class of nonlinear systems to the proposed nonlinear canonical form. And then a reduced order observer was proposed.The proposed normal forms are more generic since they contain a redundant dynamic ϑ, which in return enables to design a more robust observer.

REFERENCES

R. Abraham and J.E. Marsden. Foundation of mechanics. *Princeton, New Jersey*, 1966.

J. Back, K. Yu, and J. Seo. Dynamic observer error linearization. *Automatica*, 42:2195–2200, 2005.

D. Bestle and M. Zeitz. Canonical form observer design for nonlinear time variable systems. *International Journal of Control*, 38:419–431, 1983.

D. Boutat, A. Benali, H. Hammouri, and K. Busawon. New algorithm for observer error linearization with a diffeomorphism on the outputs. *Automatica*, 45:2187–2193, 2009.

K. Busawon, M. Farza, and H. Hammouri. A simple observer for a class of nonlinear systems. *Applied Mathematics Letters*, 11(3):27–31, 1998.

M. Darouach. Existence and design of functional observers for linear systems. *IEEE Transactions on Automatic Control*, 45(5):940–943, 2000.

M. Fliess. Analyse non standard du bruit. *Comptes Rendus de l'Académie des Sciences - Series I*, 342(10):797–802, 2006.

M. Fliess and H. Sira-Ramirez. Reconstructeurs d'état. *Comptes Rendus de l'Académie des Sciences - Series I*, 338(1):91–96, 2004.

J.-P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems-application to bioreactors. *IEEE Transactions on Automatic Control*, 37(6): 875–880, 1985.

A. Glumineau, C. Moog, and F. Plestan. New algebro-geometric conditions for linearization by input-output injection. *IEEE Transaction on Automatic Control*, 41: 598–603, 1996.

H. Hammouri and J.-P. Gautier. Bilinearization up to the output injection. *System & Control Letters*, 11:139–149, 1988.

D. Karagiannis, D. Carnevale, and A. Astolfi. Invariant manifold based reduced-order observer design for nonlinear system. *IEEE Transactions on Automatic Control*, 53(11):2602–2614, 2008.

A. Krener and A. Isidori. Linearization by output injection and nonlinear observer. *Systems & Control Letters*, 3: 47–52, 1983.

A. Krener and W. Respondek. Nonlinear observer with linearizable error dynamics. *SIAM Journal of Control and Optimization*, 30:197–216, 1985.

D. Luenberger. An introduction to observers. *IEEE Transactions on Automatic Control*, 16(6):596–602, 1971.

R. Marino and P. Tomei. Nonlinear control design: geometric, adaptive and robust. *Prentice Hall*, 1995.

H. Nijmeijer and I.M.Y. Mareels. An observer looks at synchronization. *IEEE Transactions on Circuits and Systems-1: Fundamental theory and Applications*, 44 (10):882–891, 1997.

W. Respondek, A. Pogromsky, and H. Nijmeijer. Time scaling for observer design with linearization error dynamics. *Automatica*, 40(2):277–285, 2004.

X. Xia and W. Gao. Nonlinear observer with linearizable error dynamics. *SIAM Journal of Control and Optimization*, 27:199–216, 1989.

M. Xu. Reduced-order observer design for one-sided lipschitz nonlinear systems. *IMA Journal of Mathematical Control and Information*, 26(3):299–317, 2009.

G. Zheng, D. Boutat, and J.-P. Barbot. Signal output dependent observability normal form. *SIAM Journal of Control and Optimization*, 46(6):2242–2255, 2007.

G. Zheng, D. Boutat, and J.-P. Barbot. Multi-output dependent observability normal form. *Journal of Nonlinear Analysis Series A: Theory, Methods & Applications*, 70(1):404–418, 2009.